PERCEPTION AND SHORT TERM MEMORY UNDER WORK LOAD STRESS.

MASSACHUSETTS UNIV AMHERST

11 JUN 1964

FOREWORD

As part of a program of study directed to improving human-data reception, processing and storage, this activity sponsored experimental research to investigate three variables: the strategies employed by human subjects; the organization of the information being communicated; and the characteristics of the display itself.

This report, one of three stemming from Contract N61339-1303, deals with the effect that systematically increasing the rate of input information has on short-term memory. It was found that relatively high levels of such input-load stress did not lead to an expected breakdown in the critical memory task.

The findings suggest that operators "adapt" to the high-input rates by encoding the data, i.e., imposing their own unique organization on the incoming material. This "filtering" technique permits assimilation and processing of large amounts of unconnected, meaningless material.

If the widely-held assumption that a high-enough information-input rate must disrupt performance is indeed open to question, then at least one implication of the present results seems fairly obvious: in the training for and design of high speed, man-machine closed-loop systems, greater attention should be focused on problems of storage and output processing (rather than input rate).

We have suggested some directions towards additional research to further define critical factors and test hypotheses indicated by our findings.

The two other studies performed under this contract are NAVTRADEVCEN Technical Report 1303-1 dated 11 June 1964 (The Effect of Various Modes of Rehearsal on Short-Term Recall); and NAVTRADEVCEN Technical Report 1303-3 dated 11 June 1964 (Visual After-Images as a Source of Information).

MARSHALL J. FARR
Project Psychologist
U. S. Naval Training Device Center

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

SECTION I

INTRODUCTION

The low capacity of the human operator to handle information places a limit on man-machine systems involving rapidly changing displays. As a means for increasing system performance there is currently a great deal of research attention being given to study of the factors which impose human information handling limitations. Chief among the factors assumed to be involved are those which determine the operator's reception of incoming data and those which determine his immediate memory of the data received. An understanding of the nature of these processes seems necessary as a basis for design and training principles. Such principles are especially applicable to high speed equipment characteristic of complex military systems.

Combat information centers and air traffic control centers are examples of systems in which a variety of rapidly changing visual displays are involved. Such systems include reception of raw radar returns and the transformation of these returns to aloha-numeric data presented on a secondary display. If the operator at either position is overloaded in the sense that he cannot keep up with the data, the system's performance is jeopardized. The approach of this investigation assumed that if the rate at which data are coming in is so high as to produce an actual breakdown in performance, better perrormance should result from an artificial reduction in the rate. In different terms an operator who is critically overloaded when he is given all of the data, should perform better with less than that much data. Clearly, the amount of information can be reduced to such a low level that the operator has no basis for effective performance. Thus, it can be expected, that there is some optimum input rate of information which lies between too little and too much data in a given period of time. In the examples given it is possible to regulate the information flow rate between the primary or sensing end of the system and the secondary display. There is also the possibility of storing mechanisms which can be used to regulate the rate without loss of the data.

The limitations of man's ability to process information have been ascribed to limitations in perceptual span and limitations in the span of short-term memory (Miller, 1956; Woodworth & Schlosberg, 1954). Recently this two-process distinction has been subjected to empirical investigation and is now expressable in terms of experimental operations (Teichner, Reilly & Sadler, 1961; Teichner & Sadler, 1962; Teichner & Myers, 1962; Sperling, 1960). The results of these investigations strongly suggest that short-term memory is more critical than perception to human information processing. Therefore, the present investigation was directed toward the further study of memory limiting conditions. As such it was one of three investigations carried out which were intended to evaluate hypotheses about methods for improving human data-handling (cf., Teichner & Wagner, 1964; Lewis & Teichner, 1964).

There appear to be three logically distinct, though not necessarily mutually exclusive, classes of conditions which might limit short-term memory: (a) Input processing stress which occurs, among other

situations, when the amount of information presented in a given time exceeds the maximal memory acceptance rate; (b) <u>Output processing stress</u> which occurs when the rate at which items are elicited exceeds the rate of retrieval from storage, or when the responding activity itself interferes with correct reporting; and (c) <u>Internal processing stress</u> which occurs as a result of computational and logical operations performed with items in the storage. The present experiments were concerned primarily with input stress conditions, although some problems of processing and of output stress were also considered.

Studies utilizing information measures (Klemmer & Muller, 1953; Quastler, 1955; Alluisi, Muller & Fitts, 1957; Anderson & Fitts, 1958) have demonstrated that beyond some input information rate, the information transmitted through S decreases with further increases in rate. It is important to recognize that information transmitted is a relative rather than absolute measure. That is, it is the proportion of the output associated with the input. Thus, the "breakdown" phenomenon that has been demonstrated is a breakdown of relative performance. When the percent of correctly reported items, another relative measure, is used, relative performance is perfect at low rates and decreases systematically with higher rates (e.g., Teichner, Reilly & Sadler, 1961). Thus, in the sense that increasing rates lead to a reduction in the level of performance, relative measures reflect a breakdown in performance under input stress. However, if an absolute measure of the performance level is used, such as the number of correctly reported items, the data available so far suggest that performance reaches an asymptote; there are no data yet available which suggest that, at some high level of input stress, absolute performance measures show break down. Yet common experience suggests that such a phenomenon probably exists.

The primary purpose of these studies was to explore the question of whether a performance which is critically dependent upon short-term memory breaks down with input rate increases when performance is measured in absolute terms. Since, for any fixed input time, increases in quantity of information represent increases in the input rate, a breakdown, if demonstrated, can be thought of as the result of exceeding the memory storage rate. Thus, it could be asked, given such a breakdown, whether the absolute performance level might not be increased by reducing the informational input load. In simpler terms, will S report more correct items when there is less to be recalled than when there is too much to be recalled? The studies were set up with these questions in mind.

SECTION II

EXPERIMENTAL STUDIES

Experiment I

Problem

In the studies by Teichner (1961, 1962a, 1962b) the amount of information presented to S was varied by presenting from four to nine randomly positioned letters per display and the rate was varied by exposing each display for from 0.5 to 7.0 sec. Thus, rate of input varied from 4/7 letters per sec. to 9/0.5 letters per sec. The displays were presented one-at-a-time and Ss reported in the same manner. Under these conditions for exposures of 1-sec. or more, an increase in the number of letters displayed led to a slight and negatively accelerated increase in the absolute number of letters correctly recalled.

In the present experiment an attempt was made to increase the input stress by presenting Ss with sequences of displays. After the last display in a sequence Ss were required to report the letters remembered from each display in the sequence. This task not only increased the number of letters which were to be recalled, but also required S to remember which letters were on which particular display. It was expected that this increased level of input stress would lead to "breakdown".

Sperling (1960) and Anderson (1960) have demonstrated the utility of a technique, known as partial reporting. Instead of calling for a report of the full contents of memory, a report of only a part of memory may be required. In general, more can be reported for a particular part than can be reported for that part when it is included in a full report. The advantage of partial reporting is attributed to the reduction of the interfering effect associated with the act of reporting. In our terminology this translates to a reduction in output processing stress. Thus, utilizing partial reporting as well as full reporting permits examination of the effects on short term memory of input stress at two levels of output stress. This was the general plan of this experiment. It was assumed that either reporting condition involved equivalent internal processing stress since, for any given report, it was the total contents of a single display which was demanded.

Apparatus

One hundred and sixty negative stimulus slides were made by photographing capital letters typed on 3 x 5 in. cards. The locations of letters on each slide were determined by independent random assignments of letters to the cells of a 10 x 10 matrix. Each cell was separated from the next by one typewriter space, both horizontally and vertically. The particular letters on each slide were chosen as independent random samples from the total alphabet. The number of letters on a slide, defined as the "load", was 4, 5, 6 or 7. Forty slides were made for each load. Instruction slides, telling S on which stimulus slide to report, were made by photographing black decals

3

against a white background. They read either "5th", "4th", "3rd", "2nd"
or "1st". The slides were projected via a Kodak Carousel automatic
slide projector to a screen 12.5 ft. from S's eyes. The projection
area was 64 x 44 in.; the letters were 1.5 in. high, and instruction
characters were slightly more than 3 in. high.

The exposure of a slide was regulated by a solenoid-operated
shutter mounted immediately in front of the lens of the projector. The
changing of slides, exposure time, and interslide interval were con-
trolled by a prepunched paper tape copy'ed with a two sec. pulser. The
experimental room was in semi-darkness. The Ss sat at an IBM cardpunch
and "keyed" their responses to each individual slide on separate IBM
cards. The prepunched paper tape also controlled the IBM keypunch,
automatically changing cards, and automatically locking and unlocking
the keyboard, thus controlling the time allowed for a report.

Subjects

Six volunteer undergraduate students at the University of Massa-
chusetts were used as Ss. All six had demonstrated an ability of 35
net wpm on a standard typewriter and all had served in at least one
similar type of experiment in the past. They were paid by the hour
and, in addition, competed for prizes of ten and five dollars for the
highest scores summed over all experimental conditions.

Procedure

The 40 slides of each load level were divided into eight 5-slide
sequences. Each slide was exposed for 2 sec., with 2 sec. between
slides. Two seconds after a sequence of 5 slides an instruction slide
was presented which told S on which slide he was to report. For the
Full Report (FR) condition this slide would say "5th", indicating to
S that he was to report on the last slide seen. This slide was ex-
posed for 2 sec. following which the keyboard was unlocked for 4 sec.
to allow for the keying of the report. Immediately after the report-
ing interval the next instruction slide ("4th") appeared. The sequence
was repeated until the end of the reporting interval following the last
instruction slide (which said, "1st"). A 6 sec. waiting period
followed the last reporting period before a new stimulus sequence
started. For the Partial Report (PR) condition there was only one
instruction slide and one reporting period prior to the 6 sec. wait-
ing period. In this case the instruction slide indicated that S
should report on the fourth, third, or second slide seen in the 5-slide
sequence. He was never asked to report on the first or last. The
combination of a sequence of five slides plus one or five reports will
be referred to as a trial.

With respect to report, there were four kinds of trial, namely:
(1) FR; (2) PR calling for slide 2; (3) PR calling for slide 3; and
(4) PR calling for slide 4. Each of these four kinds of trial was
called for twice under each load condition, thus yielding eight trials
for each load. One of the eight 5-slide sequences was assigned to each
of these eight trials; thus there were eight unique slide-report com-
binations. This was done for all four loads yielding a total of 32
unique slide-report combinations. These 32 combinations made up an

4

experimental session. There were five such sessions. The order of
the slide-report combinations was randomized within each load and in-
dependently randomized for each experimental session. Each of these
randomized orders was the same for all S̲s. There were 5-minute resting
periods between load treatments within a session and at least 24 but
not more than 48 hours between sessions.

Subjects were randomly assigned to two groups of three S̲s each.
The two groups were run through the five experimental sessions as
groups of three whenever scheduling permitted. During all five sessions,
one group was given the 4-load treatments in an ascending order, i.e.,
loads of 4, 5, 6 and 7 in that order. The second group of three S̲s was
given the load treatments in the opposite (descending) order. Following
this procedure, the sessions were comparable within groups for learning
or practice effects and between groups for an order effect.

Preliminary Training

Before the main experiment all S̲s were given a practice session
with a set of 1-letter slides, and then a second session with 2-letter
slides. During these practice sessions S̲s are presented with 5-slide
sequences and the same instruction slides they were later to receive
in the main experiment. These sessions were as long as was required
for S̲s to attain near-perfect performance, or for 40 trials which
ever was longer. As it turned out all S̲s reached criterion levels by
the end of the 40 sequences in both sessions.

Results

In examining the effect of order-of-load treatment, it was found
that the S̲s in the group receiving the loads in ascending order
correctly reported a fraction of a letter per slide more than the other
group. This difference showed no consistent change with load, report,
stimulus-slide position in the 5-slide sequence, or sessions. Conse-
quently, data for the two groups were combined, and all analyses are
based on the combined data.

Since the PR condition applied only to the middle three slides of
the 5-slide sequences, three separate analyses of variance were per-
formed on the data: (1) the number of letters correctly reported per
slide in the FR condition; (2) the number of letters correctly reported
per slide in the PR condition; and (3) the number of letters correctly
reported on the middle three slides of the FR condition as compared to
the number of letters correctly reported on the same three slide
positions in the PR condition. These analyses are summarized in Tables
1, 2 and 3 respectively (see Appendix A).

Full Report

The main effect of load was not significant. The main effect of
position in the 5-slide sequence and the Load x Position interaction
were both significant ($p < .01$ in both cases). The solid lines in
Fig. 1 depict these effects. The data are plotted in terms of mean
number of letters correct per slide as a function of load with a
separate graph for each of the five slide positions. Note that if the

data were collapsed across loads, the main effect of position would be a systematic decrease in performance from the first to the fourth report followed by a slight increase on the fifth report relative to the fourth. If the fifth report reversal can be accepted, the over-all trend is similar to a bowed serial position curve, i.e., a curve which is a U-function of the position of the item to-be-recalled in the series.

When considering the Load x Position interaction it may be seen that performance on the last slide seen increased as load increased, but performance on the other four slides showed no consistent change with load. This interpretation of the Load x Position interaction is further supported by examination of the data for individual $\underline{S}$s. All six $\underline{S}$s showed increasing performance with increasing load for the fifth slide seen, but there was no consistency for the other four slides.

The data shown in Fig. 1 are averaged over sessions. When plotted for individual sessions in the same manner, no orderly change could be observed corresponding to the significant Load x Position x Session interaction.

In Fig. 1 the straight dashed lines with 45° slope represent the maximum possible number of letters correct per slide. Taking the difference between the perfect performance line and the data points as a rough approximation to the inverse of the relative performance, it can be seen that, for any position in the sequence, the relative performance generally decreased as load increased. A plot of per cent correct showed the same general picture.

Partial Report

The main effects of load, position and sessions in the PR condition were all significant ($p < .01$, $p < .05$ and $p < .05$ respectively). The effects of load and position, however, require closer examination due to the significance of the Load x Position interaction ($p < .01$). The dotted lines in Fig. 1 show the effects of load and position on the mean number of letters reported correctly.

Considering only the main effect of position, it may be seen in Fig. 1 that performance first decreased slightly and then increased going from the most recently-seen slide (fourth slide presented) to the most remotely-seen slide (second slide presented). Considering only the main effect of load, it may be seen that performance first increased and then decreased as load was increased. However, this reversing load effect only characterized the reports to the second and fourth slides seen. On the third slide accuracy of report was only an increasing function of load.

A plot of the per cent correct per slide (not shown) indicates that, for all three slides, performance first increased, and then decreased as load was increased from four to seven letters per slide. It is not at all clear why the per cent correct for a load of four was lower than the per cent correct for a load of five.

Partial vs. Full Report

In the statistical analysis of performance for FR vs. PR (for the second, third and fourth slides) the data were averaged over sessions. The pairs of data lines for slides 2, 3 and 4 in Fig. 1 illustrates the comparisons made in this analysis. The main effects of load and report were both significant at the .01 level but the main effect of position was not significant. The significant effect of report is depicted in Fig. 1 where, in general, performance in the PR condition was superior to that in the FR. All effects are, however, confounded by significant interactions, namely, Report x Position, Report x Load, and Report x Load x Position.

The significant effects of experimental session (Tables 1 and 2) are illustrated in the solid lines of Fig. 2. Mean number correct per slide is plotted as a function of sessions. In general, it may be seen that performance increased gradually over the five sessions. It should be noted that Ss saw and reported on the identical slides over all five sessions. Upon questioning after the last session, all Ss reported that they had learned what to expect on some of the slides and had, in fact, recoded the letters on some of the slides into approximations to syllables, words or phrases.

Experiment II

Problem

For the FR condition the conditions of Exp. I produced no breakdown in performance with increasing loads although the input stress level had been expected to be adequate for this. Experiment II was designed to increase the stress to a still greater level in the hope of achieving a performance breakdown. In anticipation of the breakdown, Experiment II was also designed to test the hypothesis that a reduction in input information below those levels which produce breakdown will lead to a net increment in the number of correctly reported letters. The removal of some of the input information was accomplished by substituting blank slides for some of the slides in the sequence.

Apparatus

Two-hundred and eighty new slides were constructed using the procedures of Exp. I. The number of letters on a slide (load) was varied from six to nine; 50 slides were made with loads of six and 50 with loads of seven; 90 slides were made with loads of eight and 90 with loads of nine. By including the 40 slides each with loads of six and seven from Exp. I, there were a total of 90 slides for each of the four load treatments available. In addition 120 blank negative slides were made by photographing a plain white background. When projected to the screen these blank slides produced a uniformly dark gray area which covered the usual total projected area. The apparatus used for projecting the slides and for recording responses was identical to that used in Exp. I except that the prepunched paper tape was coupled with either a 1-sec. or a 2-sec. electric pulser.

Subjects

The six Ss used in Exp. I were used again. They were paid by the hour and, in addition, competed for new cash awards for the greatest number correct over all experimental conditions. In this experiment Ss were run individually through all experimental conditions.

Procedure

Only the FR condition was used. As before a trial was a full sequence of five slides plus report. Similarly the order of report was the reverse of the order of presentation. There were 15 trials per session for each of the four loads. All 15 trials at a given load level were completed before a new load level was introduced. The four load levels were always presented in ascending order of magnitude. A session contained 60 trials, 15 at each of the four levels of load. There were 6 sec. between successive trials within a given load treatment and 10 min. between load treatments within a session. There was from 24 hr. to 72 hr. between sessions.

The particular slides used for the 15 trials of any given load treatment were determined by randomly sampling 75 slides from the 90 slides available. Different random samples of slides were used for each S within a session and new random samples were used for each session.

Blank slides were inserted only in the second, third or fourth positions of the 5-slide sequence. The number of blanks inserted was either 0, 1 or 2. The position of the blank slides within a sequence was varied randomly over the 15 trials of each load level with the restriction that a blank appear equally often in the three possible positions. A different random order of blank position was used for each load level, for each S, and for each session.

Two exposure times (1 or 2 sec.) and three numbers of blank slides (0, 1 or 2) were combined orthogonally to produce six different combinations of treatment. A given combination, fo. a given S was constant over each experimental session. The treatments were administered to the S in a balanced Latin square which is described in Table 4.

Two sec. after the last slide of any sequence, five 4-sec. time periods were made available for reporting on each of the five slides in the sequence. Even though no instruction slides were used in this experiment there was 2 sec. between adjacent slide reporting intervals. The Ss were instructed to use each of the five periods for the appropriate slide reports, but if the slide to be reported during a given period were a blank slide, they were to wait out the time allotted for the report without keying the cardpunch.

Results

For reasons beyond our control only four Ss completed all six sessions (see Table 4). It was necessary, therefore, to use only these Ss and to analyze the data under the assumption that practice effects and first order sequential effects were not significant. It should be

noted that, if the assumption made above were in fact not true, the practice and the sequential order effects were confounded with the other experimental variables.

An analysis of variance (Table 5) was performed on the total number of letters correctly recalled per 5-slide sequence. Blank slides, of course, contributed nothing to the total. The effect of exposure time was not significant. The effects of load, number of blanks, and the Load x Blanks interaction were significant, each at $p < .01$.

Figure 3 presents the data in terms of the mean number correct per slide in the 5-slide sequence as a function of load with blanks and exposure time as parameters. The figure shows that under all conditions the number correct increased as the load increased. Also, for each level of load, an increase in the total number of slides which actually contained letters led to an increase in the number correct. Figure 3 also suggests that performance in the zero blank condition increased somewhat more rapidly with increasing load than in either the 1- or 2-blank conditions. Thus, it is seen that the significant Load x Blank interaction is orderly.

Comparisons were also made between comparable conditions of this experiment and Exp. I. It was found that for the four $\underline{Ss}$ common to both experiments, the Exp. I mean was 2.76 correct per slide, while for the present experiment it was 3.29 letters correct. Reports by the $\underline{Ss}$ after the last experimental session indicate that a considerable portion of this continued improvement with practice might be attributed to an increased frequency of encoding the letters on each slide into more easily remembered approximations to syllables, words, and phrases.

Experiment III

Problem

Input loads as high as nine letters per slide in a 5-slide sequence failed to produce breakdown, and the introduction of blank slides failed to increase (actually decreased) the reportable contents of short-term memory. Experiment III increased input load stress still further by increasing the number of slides in the slide sequence. In addition, since the $\underline{Ss}$ of Exps. I and II were very familiar with the slides, Exp. III employed naive $\underline{Ss}$.

Apparatus

Eighty-one slides from each of loads 6, 7, 8 and 9 were randomly chosen for use in this experiment. In addition to these, blank slides were also used.

The projection and recording apparatus was identical to that described previously.

Subjects

Thirty volunteer students on the University of Massachusetts campus during the summer session were used in this experiment. All of

these $\underline{S}$s demonstrated a minimum typing speed of 35 wpm and were all at least 18 years of age. All $\underline{S}$s were paid by the hour and in addition competed within their respective groups for cash awards for the highest score summed over all experimental conditions.

Procedures

The 30 $\underline{S}$s were divided into three groups of 10 each. Each group was randomly assigned to a different length of sequence condition. The three length conditions were 5, 7, or 9 slides per sequence. These sequences of slides were followed respectively by sequences of 5, 7 or 9 slide-report periods, each slide to be reported in reverse of the order of presentation. A trial was defined as the combination of a sequence of slides and the corresponding reports.

Each $\underline{S}$ responded for 36 trials in three blocks of 12 trials each, corresponding to the three blank slide conditions. The order of administering the blank slide conditions was partially balanced over $\underline{S}$s. Each block of 12 trials consisted of four sets of three trials each for loads 6, 7, 8 and 9 (sets administered in that order). Each set of three trials consisted of one trial each for each of the three possible positions of blank slides. The load and blank slide conditions were as defined in Exp. II. Blank slides were always inserted into either one or two of the middle three positions of a given sequence. For example, in a sequence of length 7, blank slides, when employed, could be inserted into positions 3, 4 and/or 5. Within any blank condition all three possible combinations of positions were employed but the order of presentation of the positions over the set of three trials was independently randomized for each load and each length of sequence. The actual slides, and the order of the slides used in a given set of three trials were the same for all $\underline{S}$s in a group. Blank slides were used temporarily to replace the letter slides when required.

Slides were exposed for 1 sec. and there was 1 sec. between successive slides in a given sequence of slides. Two sec. after the exposure of the last slide in a sequence the IBM keyboard was unlocked for 4 sec. to allow for the first report. Thereafter, for the duration of the sequence of reports, the keyboard was unlocked for 4 sec. every 5 sec. Thus, there were 1-sec. intervals between successive reports. No instruction slides were used.

The $\underline{S}$s were run individually through all appropriate conditions in one day's session. There were rest periods of about 2 min. between successive load treatments within a given blank condition and about 10 min. between successive blank conditions.

Preliminary Training

Every $\underline{S}$ was put through three separate practice sessions before participating in the main experiment. The $\underline{S}$s were paid by the hour during these sessions also. Details of these sessions are to be found in Appendix C. Each of the first two sessions involved practice on 80 to 100 single slides of loads 4 and 5. The third session involved 60 slides with loads of 4 through 9. Just prior to the first session of the main experiment $\underline{S}$s were given six trials of the sequence length

appropriate for the group to which they had been assigned. They reported at the end of the sequence, as they did in the main experiment. There were three trials each of loads 4 and 5. Load was constant within trials. Ten min. after this practice session the main experiment was started.

Results

The results of this experiment were analyzed in terms of the total number correct summed across the three trials which made up a Length x Blank x Load condition. The analysis of variance is summarized in Table 6.

This analysis showed the main effects of length, load, and blanks all to be significant ($p < 0.01$ in each case). Figure 4 illustrates the differences involved in these effects. The dependent variable in this figure is the total number correct over the same three trials used in the analysis of variance summed over the 10 $\underline{S}$s in each group (length of sequence). It may be seen that as the length of sequences increases, there is an over all tendency for total number correct to increase. The effect of an increase in load leads, also, to an overall increase in the total number correct and this is true for each of the length treatments. The depicted tendency for load to have a greater effect for the sequences of length 9 than for the shorter sequences is supported by the significant Length x Load interaction ($p < 0.01$).

The significant effect of blanks shows up in each of the length conditions but is not as consistent as the other two main effects. In general, as number of blank slides in a sequence increased the total number correct decreased. The load by blank interaction, which was also significant ($p < 0.01$), can also be seen for each of the length conditions, though it is not very orderly.

SECTION III

DISCUSSION

For both experienced and naive $\underline{S}$s it was found that increasing
the length of a series of slides, increasing the number of information
slides per unit time, and increasing the number of letters per slide
all lead to an increase, or at least no decrease, in the total number
of letters correctly recalled when the $\underline{S}$ is asked to make a full re-
port (FR) of what he was shown. For these conditions it appears rea-
sonable to conclude that increases in input stress do not lead to
breakdown.

For the partial report (FR) condition, however, performance
appears to break down for certain slide positions in a sequence of
slides. According to the usual interpretation (Anderson, 1960; Sperl-
ing, 1960) of the PR condition, this condition should have resulted
in a reduction in output processing stress relative to the FR condi-
tion, with all else remaining approximately constant. This leads to
the uncomfortable conclusion that reducing output stress leads to
conditions which will yield break down in performance for high levels
of input stress. A re-examination of the situation, however, suggests
the following hypothesis: PR conditions such as those of Experiment
I impose processing and output stress which are absent in the FR con-
dition. The FR condition requires $\underline{S}$ to recall each slide in sequence
and he has 4 sec. to report on each of these slides. The PR condition,
on the other hand, requires that $\underline{S}$ scan his memory until he finds the
appropriate slide after which he must make a report, but he has only
a total of 4 sec. to accomplish both the scan and the report. As
the number of letters per slide is increased the differences between
slides is reduced, and the selective retrieval of just one becomes
increasingly difficult. Further, if $\underline{S}$ must scan through the slides
item by item, it necessarily takes longer to do it if there are more
items. The lack of breakdown for the central slide position, how-
ever, remains to be explained. Additional experimental exploration
is needed.

The ineffectiveness of increases in input load stress to produce
a decrement in the amount stored in short term memory remains to be
explained. Three ideas are relevant to any possible explanation(s).
First, what would a random guessing machine do if it emitted as many
responses (guesses) as did the $\underline{S}$s in the experiments? Second, to
what extent did the $\underline{S}$s encode ("chunk") the letters on each slide?
And, third, do the $\underline{S}$s impose their own input filter, essentially
ignoring all information that they feel they cannot handle?

1. The expected number correct for any given load and for any
given number of responses may be calculated via the hypergeometric
distribution (Feller, 1957). Figure 5 depicts the comparison of
actual number correct with the number expected correct by chance for
certain comparable conditions across the three experiments. The four
experienced $\underline{S}$s of Experiments I and II tend to emit at least as many
responses per slide as there were letters per slide, emitting con-
sistently more in Experiment II. The naive $\underline{S}$s of Experiment III emit
fewer responses than there were letters, especially for the higher

12

loads. These differences in number of responses account for the differences in the expected numbers correct. In general, performance is above chance for all conditions, but approaches towards chance as load is increased. The data cannot be accounted for simply in terms of a random guessing machine. The $\underline{S}$s actually store something in short-term memory.

2. According to the casual verbal reports of the $\underline{S}$s during and after the experiments, the something that is being stored may very frequently be an encoding, or "chunking" (Miller, 1956) of the letters on a slide, and not just a sequence of randomly juxtaposed letters. The encoding may take the form of an approximation to a syllable, word, or even a phrase. The experienced $\underline{S}$ reported this encoding activity more consistently than did the naive $\underline{S}$s, and, in general, the more encoding reported, the better $\underline{S}$ appeared to do. It would appear, therefore, that any account of the data should consider such encoding activity. The "encodability" of the random letters on a slide, for example, may be expected to increase as the number of letters on a slide increases (there is a greater chance of finding a high frequency sequence of letters), and this effect is counter to that expected from load stress.

3. The constant, or near constant (slightly increasing), number of letters correct as load increased suggests that $\underline{S}$s filter the input they are exposed to and accept, on the average, only as much as they can handle. The amount that the filter is set for may be a function of the amount that can be "chunked", and this, in turn, a function of both the load and the experience of the $\underline{S}$s. Unfortunately, the concept is theoretically useless for accounting for the present data unless some independent measure can be obtained of the number of letters for which the filter was set for the various experimental conditions.

SECTION IV

CONCLUSIONS AND RECOMMENDATIONS

The results suggest that:

1. Increases in input stress do not produce breakdown in performance on a short term memory task.

2. Decreasing input information decreases the amount stored in short term memory.

Consideration of these results lead to the following important hypotheses:

1. Breakdown in performance on a short-term memory task must involve internal processing or output processing stresses.

2. Prediction of recall, for any but the simplest of tasks, requires an understanding of the encoding and filtering activities of the S_s.

It is generally assumed by those responsible for training and human engineering in the design of high speed, man-operated systems that the operator may be so overloaded as a result of the speed and amount of input that his performance will suffer and, with sufficient input rate, may break down. For example, a radar observer, whether a pilot, a receiver of alpha-numeric data in an air traffic control or combat information center, or the receiver of raw radar returns at the sensory end of such centers may receive so much data so fast that he may not only lag behind the input, he may be in danger of collapse as a receiver. The present results suggest that the input rate assumption may lead to improper emphasis in training and design, that emphasis may be put more effectively on problems of internal and output processing stress in training and design.

Before the implications of the present results can be put into practice, more must be known about the internal and output processing systems and about the means by which human operators filter and encode data. Our analysis of the present study leads to a number of hypotheses in these regards which appear to define the critical needs for further research. These can be expressed as follows:

1. In tasks involving high input rates, operators will impose their own filtering on the input, accepting only as much as they can handle consistent with task requirements.

2. Tasks involving processing load stress will lead to a breakdown in operator performance as input load stress is increased. The characteristics of critical processing load tasks should be explored.

3. Breakdown in performance may be delayed, reduced and/or eliminated by a) restructuring the task, b) training, c) appropriately filtering the input. These procedures should be studied.

4.   Input information is usually assimilated in meaningful chunks (as opposed to isolated elements). The greater the opportunity and ability to chunk, the greater the amount of information the operator can assimilate, and the more resistant he is to breakdown.

5. Whether filtering of input will reduce or eliminate breakdown will depend on whether or not chunks or isolated elements are filtered out.

It is recommended that research be initiated to test these expectations.

SECTION V

REFERENCES

Alluisi, E. A., Muller, P. F., and Fitts, P. M. An information analysis
of verbal and motor responses in a forced-paced serial task. J. exp.
Psychol., 1957, 53, 153-158.

Anderson, N. S. and Fitts, P. M. Amount of information gained during
brief exposures of numerals and colors. J. exp. Psychol., 1958,
56, 362-369.

Anderson, N. S. Poststimulus cuing in immediate memory. J. exp. Psychol.,
1960, 60, 216-221.

Feller, W. An Introduction to Probability Theory and its Applications.
John Wiley, 1957, New York. Second edition, Chapter 2.

Klemmer, E. T. and Muller, P. F. The rate of handling information. Key
pressing responses to light patterns. HFORL Memo Report No. 34.
Operations Applications Lab., Air Research and Development Command,
Bolling Air Force Base, Washington 25, D. C. March 1953.

Lewis, Mary E., Teichner, W. H. The effect of various modes of rehearsal
on short-term recall. U. S. Naval Training Device Center Tech. Rept.:
NAVTRADEVCEN 1303-1, 1963.

Miller, G. A. The magical number seven, plus or minus two: Some limita-
tions on our capacity for processing information. Psychol. Rev.,
1956, 63, 81-87.

Quastler, H. Information Theory in Psychology: Problems and Methods.
Glencoe, Ill.: Free Press, 1955, (p. 246).

Sperling, G. The information available in brief visual presentations.
Psychol. Monogr., 1960, 74 (11, Whole No. 498).

Teichner, W. H., Reilly, R., and Sadler, E. Effects of density on identi-
fication and discrimination in visual symbol perception. J. exp.
Psychol., 1961, 61, 494-500.

Teichner, W. H. and Sadler, E. Effects of exposure time and density on
visual symbol identification. J. exp. Psychol., 1962a, 63, 376-380.

Teichner, W. H. and Myers, J. L. Training aspects of decision making.
Phase II. U.S. Naval Training Device Center Tech. Rept:
NAVTRADEVCEN 588-2, June 1962b.

Teichner, W. H. and Wagner, M. Visual after-images as a source of
information. U.S. Naval Device Center Tech. Rept.: NAVTRADEVCEN
1303-3, 1963.

Woodworth, R. S. and Schlosberg, H. Experimental Psychology. New York:
Holt, 1954.

APPENDIX A

Table 1

Summary of Analysis of Variance of Number of Letters
Correct per Slide for the Full Report
in Experiment I

Source	df	MS	F
Between Subjects (S)	5		
Within Subjects	594		
Load (L)	3	17.584	3.124
Position (P)	4	317.561	26.150**
Session (Se)	4	39.690	13.555**
L x P	12	18.202	7.356**
L x Se	12	3.798	1.086
P x Se	16	1.789	0.847
L x P x S	48	4.168	1.692**
L x S	15	5.628	
P x S	20	12.144	
Se x S	20	2.928	
L x P x S	60	2.474	
L x Se x S	60	3.498	
P x Se x S	80	2.112	
L x P x Se x S	240	2.463	

** p<0.01

APPENDIX A

Table 2

Summary of Analysis of Variance of Number of Letters
Correct per Slide for the Partial Report
In Experiment I

Source	df	MS	F
Between Subjects (S)	5		
Within Subjects	354		
Load (L)	3	82.892	14.053**
Position (P)	2	57.570	7.294*
Session (Se)	4	14.802	3.887*
L x P	6	20.003	3.813**
L x Se	12	4.903	1.294
P x Se	8	1.149	0.386
L x P x Se	24	4.629	1.450
L x S	19	5.898	
P x S	10	7.893	
Se x S	20	3.808	
L x P x S	30	5.246	
L x Se x S	60	3.788	
P x Se x S	40	2.981	
L x P x Se x S	120	3.192	

*p<0.05
**p<0.01

APPENDIX A

Table 3

Summary of Analysis of Variance of Number of Letters
Correct per Slide for the Full Report vs. the
Partial Report in Experiment I

Source	df	MS	F
Between Subjects (S)	5		
Within Subjects	138		
Load (L)	3	253.987	11.088**
Report (R)	1	1400.000	47.310**
Position (P)	2	71.045	2.909
L x R	3	276.510	6.417**
L x P	6	45.587	2.781*
R x P	2	573.345	8.494**
L x R x P	6	102.338	5.275**
L x S	15	22.905	
R x S	5	29.592	
P x S	10	24.424	
L x R x S	15	43.089	
L x P x S	30	16.394	
R x P x S	10	67.498	
L x R x P x S	30	19.399	

* p<0.05
** p<0.01

APPENDIX A

Table 4

An Illustration of the Balanced Latin Square Design
Used to Control for Practice and First Order
Sequential Effects

Subject	Sessions					
	1	2	3	4	5	6
1	10*	11	22	12	21	20
2**	11	12	10	20	22	21
3	12	20	11	21	10	22
4**	20	21	12	22	11	10
5	21	22	20	10	12	11
6	22	10	21	11	20	12

*   The first digit in each cell represents the exposure time (in sec.)
and the second digit in each cell represents the number of blanks in-
serted per sequence.

**   These two $s did not complete all six experimental sessions.
Their data were subsequently discarded.

APPENDIX A

Table 5

Summary of Analysis of Variance of Total Number Correct
Summed over Five Slides in a Sequence
in Experiment II

Source	df	MS	F
Between Subjects (S)	3		
Within Subjects	92		
Load (L)	3	17039.70	32.05**
Exposure (E)	1	23343.70	4.91
Blank (B)	2	36723.85	22.73**
L x E	3	877.63	3.56
L x B	6	1310.33	8.95**
E x B	2	482.10	0.19
L x E x B	6	255.30	1.16
L x S	9	531.60	
E x S	3	4756.57	
B x S	6	1615.55	
L x E x S	9	246.21	
L x B x S	18	146.48	
E x B x S	6	2575.47	
L x E x B x S	18	219.64	

** $p < 0.01$

APPENDIX A

Table 6

Analysis of Variance for Total Number Correct Summed
Across the Three Trials Within a Blank
Condition in Experiment III

Source	df	MS	F
Between Subjects	29		
Length (le)	2	6596.61	6.154**
Subjects w. groups	27	1071.91	
Within Subjects	330		
Load (Lo)	3	1694.96	40.077**
Le x Lo	6	138.01	3.263**
Lo x S	81	42.29	
Blank (B)	2	980.18	15.372**
Le x B	4	21.61	0.339
B x S	54	63.76	
Lo x B	6	126.38	5.788**
Le x Lo x B	12	16.85	0.772
Lo x B x S	162	21.83	

** $p < 0.01$
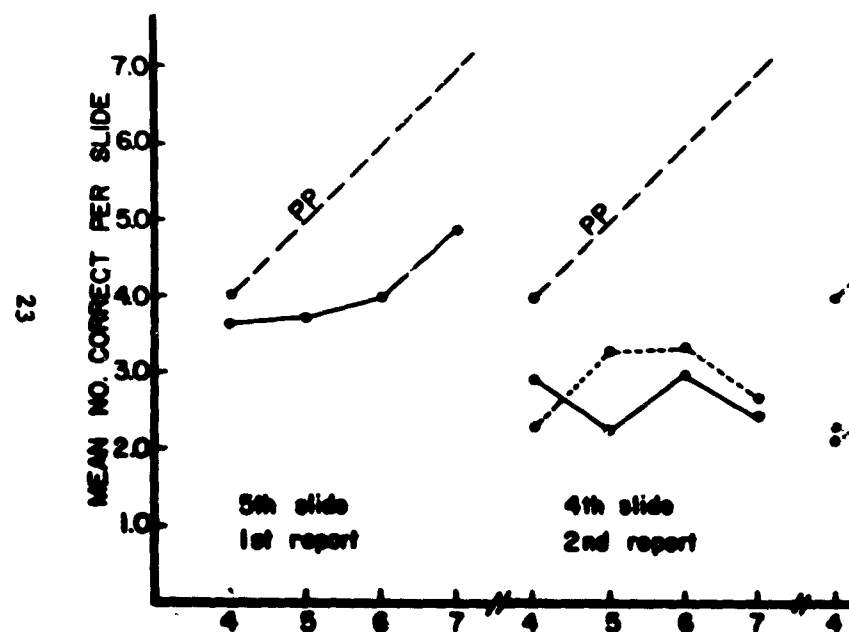
Figure I.   The mean number correct as a function of load

PP = Perfect Performance

——————— FR

----------- PR

PP

PP

2nd slide   1st slide
4th report   5th report

6   7   4   5   6   7   4   5   6   7
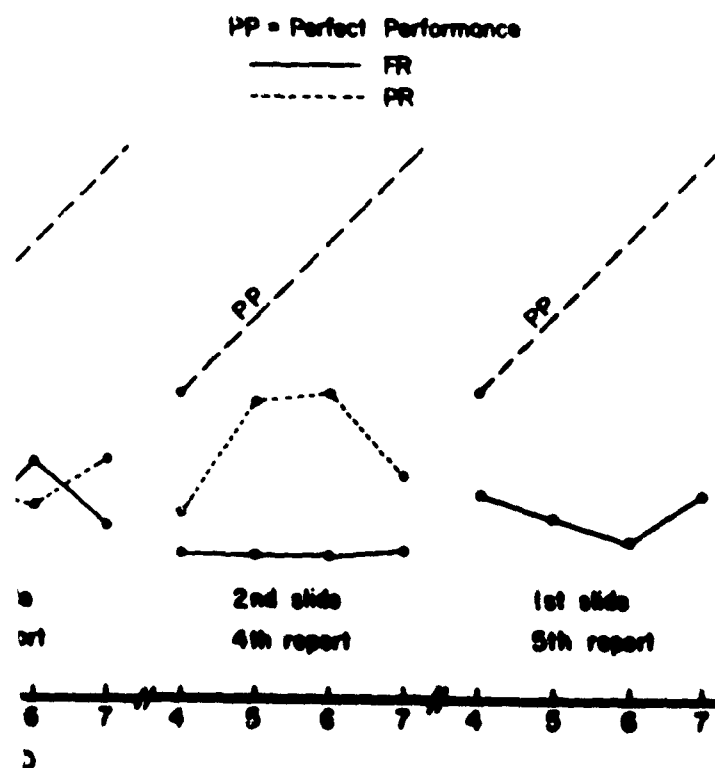
on for the full and partial report conditions in Exp. I.
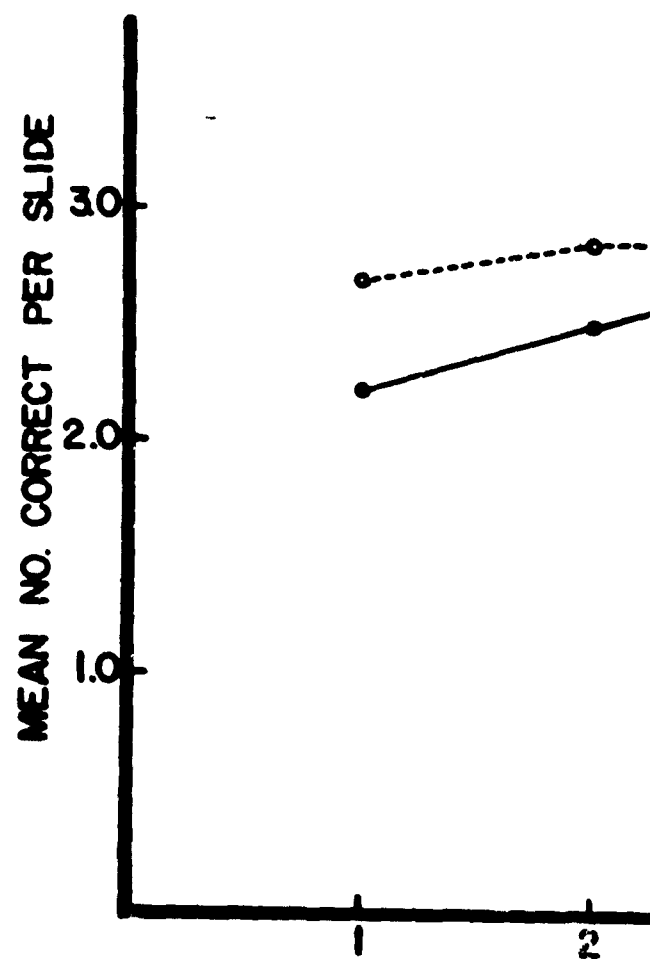
Figure 2. The mean number
for the full and partial rep

——————— FR

- - - - - - - - PR

3   4   5
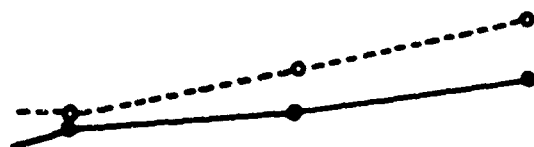
S S I O N

>ct as a function of session

onditions in Experiment I.
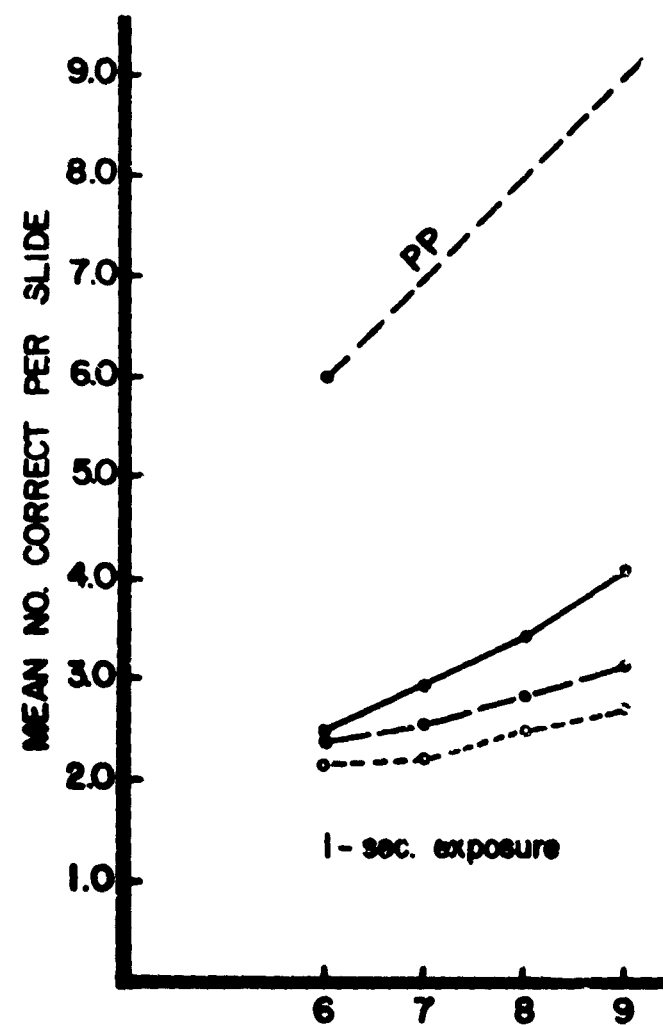
Figure 3. The mean number cor:
with blanks as a parameter for

PP

Zero Blanks

One Blank

Two Blanks

2-sec. exposure

6   7   8   9.
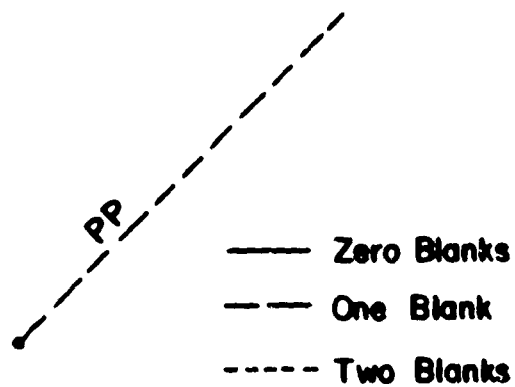
function of load and exposure

II.
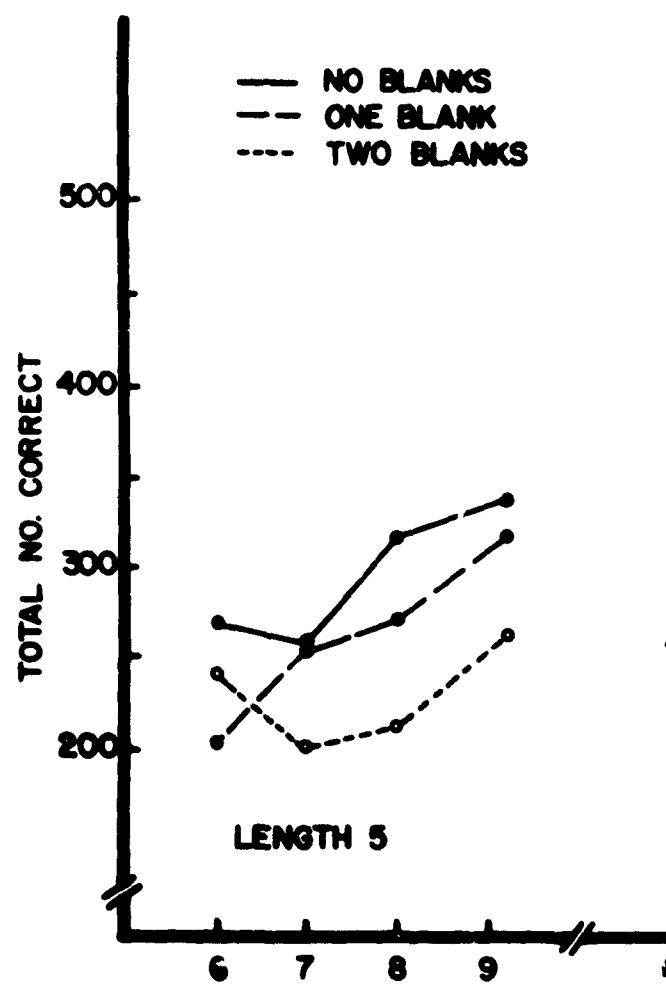
Figure 4. The total number correct c
parame
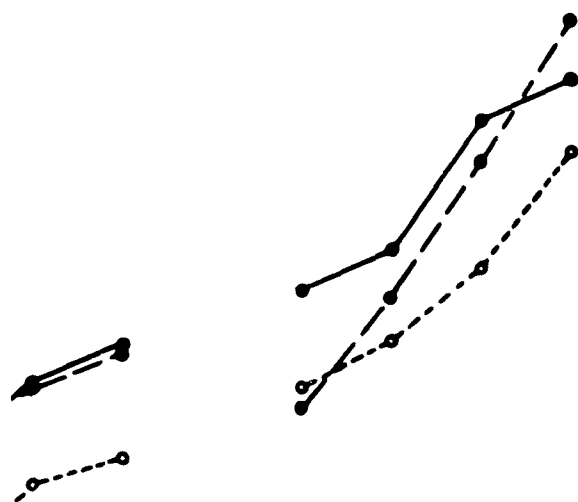
7   LENGTH 9

8   9   6   7   8   9

D

n of load and length with blanks as a
periment III.
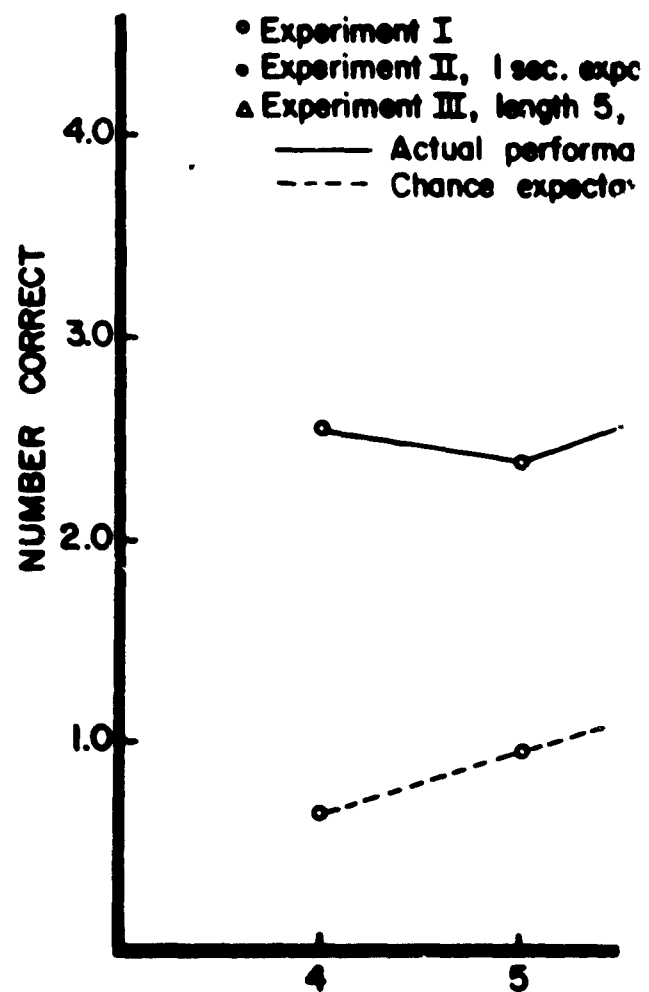
Figure 5.   A comparison of
bas

lanks
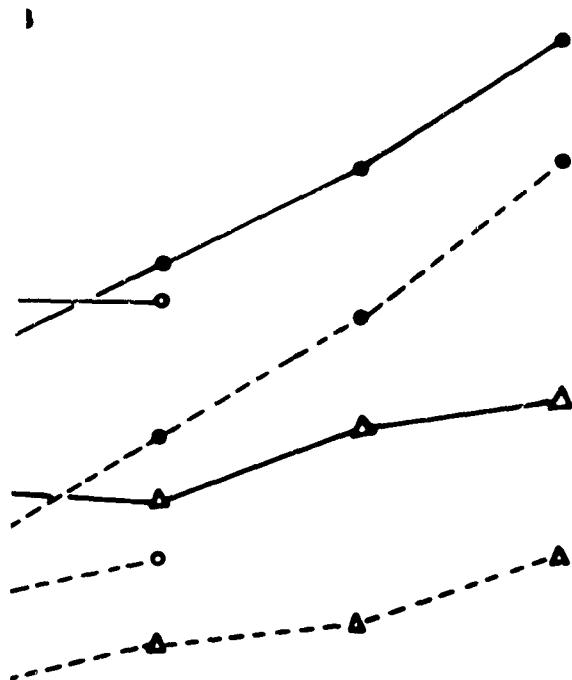
.OAD

·formance and predicted performance

once alone.

APPENDIX C

Details of Preliminary Training
for Experiment III

In the first three sessions which lasted about 3/4 hour each,
Ss were run in groups of two and/or three when ever scheduling permit-
ted. The first session was divided into two equal time periods.
During the first period Ss were shown slides of load 4. During the
second period Ss were shown slides of load 5. In both periods single
slides were exposed and then immediately reported on via the IBM
cardpunch. Slides were exposed for 2 sec. and there was 4 sec. allowed
for each slide report. The second session was identical to the first
(the order of slides within a load was randomly shuffled relative to
session) except that slides were only exposed for one second. In
each of the first two practice sessions Ss were shown and reported on
80-100 slides from each of loads 4 and 5.

The third practice session first replicated session number two
with the order of slides within a load again randomly shuffled. Then,
Ss were shown a new series of slides which consisted of five slides
each of loads 4 through 9. In this latter series of slides, the loads.
were kept in 5 slide blocks. the order of load presentation was
ascending, and slides were exposed one at a time (for 1 sec.) followed
immediately by a 4 sec. report interval. This series of 30 slides was
then immediately replicated.